

Reassessing Benzene Risks Using Internal Doses and Monte-Carlo Uncertainty Analysis

Louis Anthony Cox Jr.

Cox Associates, Denver, Colorado

Human cancer risks from benzene have been estimated from epidemiological data, with supporting evidence from animal bioassay data. This article reexamines the animal-based risk assessments using physiologically based pharmacokinetic (PBPK) models of benzene metabolism in animals and humans. Internal doses (total benzene metabolites) from oral gavage experiments in mice are well predicted by the PBPK model. Both the data and the PBPK model outputs are also well described by a simple nonlinear (Michaelis-Menten) regression model, as previously used by Bailer and Hoel (Metabolite-based internal doses used in risk assessment of benzene, *Environ Health Perspect* 82:177–184 (1989)). Refitting the multistage model family to internal doses changes the maximum-likelihood estimate (MLE) dose-response curve for mice from linear-quadratic to purely cubic, so that low-dose risk estimates are smaller than in previous risk assessments. In contrast to Bailer and Hoel's findings using interspecies dose conversion, the use of internal dose estimates for humans from a PBPK model reduces estimated human risks at low doses. Sensitivity analyses suggest that the finding of a nonlinear MLE dose-response curve at low doses is robust to changes in internal dose definitions and more consistent with epidemiological data than earlier risk models. A Monte-Carlo uncertainty analysis based on maximum-entropy probabilities and Bayesian conditioning is used to develop an entire probability distribution for the true but unknown dose-response function. This allows the probability of a positive low-dose slope to be quantified; it is about 10%. An upper 95% confidence limit on the low-dose slope of excess risk is also obtained directly from the posterior distribution and is similar to previous q_1^* values. This approach suggests that the excess risk due to benzene exposure may be nonexistent (or even negative) at sufficiently low doses. Two types of biological information about benzene effects—pharmacokinetic and hematotoxic—are examined to test the plausibility of this finding. A framework for incorporating causally relevant biological information into benzene risk assessment is introduced, and it is shown that both pharmacokinetic and hematotoxic models appear to be consistent with the hypothesis that sufficiently low concentrations of inhaled benzene do not create an excess risk. — *Environ Health Perspect* 104(Suppl 6):1413–1429 (1996)

Key words: biologically based risk assessment, dose-response modeling, Monte-Carlo uncertainty analysis, simulation modeling, causal modeling

Introduction

For risk managers and public health decision makers, the most important questions about human cancer risks from exposures to low concentrations of benzene are *a*) Is there any excess risk at low exposure concentrations (e.g., below 1 ppm)? *b*) If so, how large is the excess risk, and how does it vary with changes in exposure concentration and timing?

After more than 20 years of vigorous investigation using increasingly sophisticated techniques of molecular biology, computer and statistical risk models, and epidemiological data analyses, the answers to these questions remain unknown. This paper briefly reviews how inferences about low-dose human cancer risks are traditionally drawn from animal bioassay data. It

then suggests that a different approach is needed for benzene because benzene biology is inconsistent with the assumptions embedded in the usual approach. Three methods that may lead to more realistic and informative estimates of benzene risks are introduced as follows:

- Use of Monte-Carlo uncertainty analysis to quantify the probability distribution of the entire dose-response function for benzene-induced tumors in male B6C3F1 mice. In contrast to previous Monte-Carlo uncertainty analyses for benzene that treat uncertain model parameter values as random variables (1), here the true response probabilities in different experimental dose groups are treated as random variables.
- Use of two species-specific physiologically based pharmacokinetic (PBPK) models, one for mice and the other for humans, to quantify internal dose-response functions and to extrapolate risks from mice to humans. This approach leads to conclusions opposite to those previously reached using a PBPK model for mice and a traditional allometric scaling (2) approach to interspecies extrapolation.
- A modeling framework for incorporating additional biological knowledge—about cytotoxicity, cell kinetics, and genotoxicity, as well as the information about pharmacokinetics and metabolism contained in the PBPK model—into benzene dose-response models.

The quantal bioassay data displayed in Table 1 motivate and illustrate the methods developed. These data are for the most sensitive species, strain, and sex found among many animal bioassay studies reviewed by Crump and Allen (3). They have been referenced by the Occupational Safety and Health Administration (OSHA) (4) in its regulation of benzene.

Applied risk assessment attempts to draw inferences from data sets such as that

This paper was presented at Benzene '95: An International Conference on the Toxicity, Carcinogenesis, and Epidemiology of Benzene held 17–20 June 1995 in Piscataway, New Jersey. Manuscript received 16 January 1996; manuscript accepted 14 June 1996.

I thank the American Industrial Health Council Delivered Dose Working Group, the American Petroleum Institute, and the Western States Petroleum Association for their support of research on new methods for dealing with uncertainties in complex, biologically based risk models. I am especially grateful to P. Beatty, C. Lapin, and M. Paxton for insightful and thoughtful discussions.

Address correspondence to Dr. L.A. Cox Jr., Cox Associates, 503 Franklin Street, Denver, CO 80218. Telephone: (303) 388-1778. Fax: (303) 388-0609. E-mail: lcoxdenver@aol.com

Abbreviations used: AML, acute myelogenous leukemia; AUC, area under curve; CFU-GM, colony-forming unit, granulocyte-macrophage; CFU-S, colony-forming unit, spleen; CDF, cumulative probability distribution function; LMS, linearized multistage model; MLE, maximum likelihood estimate; PBPK, physiologically based pharmacokinetics; PDE, partial differential equation; WBC, white blood cell.

Table 1. Oral gavage data for B6C3F1 male mice.

x = Dose in mg/kg/day benzene	0	25	50	100
$n(x)$ = Number of animals exposed	50	50	50	50
$r(x)$ = Fraction of mice with tumors (squamous cell carcinomas)	0	0.08	0.40	0.74

From Crump and Allen (3), Appendix B, reproduced by the U.S. EPA (5).

described in Table 1 to answer the following types of questions about risk: *a*) What is the excess risk at 25 mg/kg/day? Although the observed risk is 0.08, this observation contains sampling variability and ignores the information in the other data points. Additional analysis is needed to obtain a more informative answer. *b*) What is the excess risk in male mice exposed to 1 mg/kg/day? Answering this question requires extrapolation beyond the range of the data, and therefore requires some sort of risk model or assumptions. *c*) What is the excess risk in humans exposed via inhalation to 1 ppm for 45 years, starting at age 20? This type of question is of most interest to risk managers. However, it cannot be answered purely on the basis of statistical analysis of the data. Instead, modeling assumptions are required to permit extrapolation across species, route of administration, and dose regimen. For applied work, it is essential to understand how robust risk conclusions are to plausible changes in the model assumptions. If slightly more realistic assumptions lead to large changes in estimated risk, which turns out to be the case for benzene, then the uncertainty about risks may be driven more by model uncertainty than by data uncertainty. However, model uncertainty is more difficult to quantify and manage. It has often been disregarded in regulatory risk assessments of benzene.

Traditional risk assessment, discussed next, infers from the data in Table 1 that there is a strong, positive dose-response relation for benzene that extends even to very low doses. However, traditional risk assessment incorporates assumptions (e.g., that risk is determined by total administered dose, or that effects of exposure on cell proliferation and cytotoxicity may be ignored) that may not be appropriate for benzene. Reanalysis of the data in Table 1 using more flexible methods indicates that there is no evidence for an increased risk at low exposure levels.

Traditional Statistical Risk Assessment of Benzene

A traditional regulatory risk assessment proceeds as follows:

Step 1. A dose metric and a statistical risk model are chosen to relate the administered dose history to tumor probability.

Step 2. The parameters of the risk model are estimated from the available experimental data, e.g., by choosing parameter values that maximize the likelihood of the data.

Step 3. Uncertainty about the parameters is quantified using statistical theory. Uncertainty about the dose metric and the statistical risk model formula are usually either ignored or treated by making risk estimates for several different choices of dose metrics and models.

Step 4. Risks are interpolated among dose levels and extrapolated beyond the conditions of the experimental data based on the chosen dose metric and risk model, e.g., to new species and new time patterns of dose administration. Specifically, the dose metric establishes which administered dose histories are considered to create equal risks, while the risk model quantifies how large these risks are.

These steps will now be illustrated for the data in Table 1.

Step 1. Selection of a Dose Metric

A dose metric converts any specified detailed administered dose history (a time sequence of hourly or daily benzene intakes over a period of weeks to years) into a single corresponding number called the "dose," typically plotted on the horizontal axis of the dose-response curve. Any two histories that are assigned the same number (dose) by the dose metric are thereby assumed to create the same risk, even if the time pattern or route of administration is different, and even if different species and strains are involved. An example of a dose metric that has been applied to the data in Table 1 (5) is mg-lifetimes of benzene per unit of surface area of the exposed animal. For any dose history (e.g., 60 days of exposure to 25 mg/kg/day), the corresponding dose number is obtained by calculating the total quantity of benzene (e.g., 1500 mg), multiplying it by the fraction of life exposed (e.g., 60 days/1000 days for a species and strain that lives 1000 days on average), and dividing by surface area (approximated as the two-thirds power of the animal's body weight). In the remainder of this article, x will denote the dose variable defined by the selected dose metric (e.g., mg-lifetimes per unit surface area).

The most common choice of dose metric for risk interpolation and extrapolation within a single species, strain, and sex is the area-under-curve (AUC) dose metric, which assigns the same dose (and hence the same risk) to any two dose histories that have the same integrated total dose (AUC for the administered dose history). For example, risk models in which equal ppm-hr of benzene or equal mg/kg of benzene create equal predicted

risks independent of the detailed time pattern of dose administration are using AUC dose metrics.

Step 2. Parameter Estimation via Maximum Likelihood

Once a dose metric has been selected, experimental data such as those in Table 1 can be plotted to obtain an empirical dose-response function. A statistical model of the dose-response relation is typically represented by a multivariate function (letters in bold type indicate vectors).

$$p = f(x, q)$$

where

p = lifetime probability of tumor

x = dose (computed via the dose metric selected in Step 1)

q = a vector of model parameters

f = assumed model formula.

For example, the data in Table 1 might be assumed to result from the familiar multistage risk model, which has the form

$$f(x, q) = 1 - \exp[-(q_0 + q_1x + q_2x^2 + \dots + q_kx^k)],$$

where

$$q = (q_0, q_1, q_2, \dots, q_k).$$

By varying q , the model dose-response function $f(x, q)$ may be made to approximate the empirical dose-response functions obtained by plotting the fraction of animals with tumors against the dose, x . The quality of the approximation is quantified by a specific criterion, such as the likelihood criterion, which quantifies the likelihood of the empirical curve, as predicted from the model curve; or the mean squared difference between the empirical and model values. Then q may be selected to optimize the selected criterion. This provides an estimate of q based on the observed experimental data, the assumed model, and the selected dose metric.

As an example, the maximum-likelihood estimate (MLE) of q for the multistage model was computed for the data in Table 1, assuming an AUC dose metric of mg/kg/day of administered benzene. The value of q that maximizes the likelihood function was found via a numerical optimization computer program assuming that only parameter vectors satisfying the constraint $q \geq 0$ are to be considered, as is

traditional in regulatory risk assessment. The MLE for q is

$$q^* = (q_0, q_1, q_2)^* = (0, 0.00145, 0.00013) \\ \text{(MLE parameter estimate).}$$

These parameter values maximize the likelihood function

$$L(q) = \text{Pr}(\text{Data} | q) \\ = \prod_x f(x, q)^{n(x)r(x)[1-f(x, q)]^{1-r(x)n(x)}},$$

where $n(x)r(x)$ denotes the number of animals that developed tumors out of the $n(x)$ animals in dose group x . They correspond to the MLE dose-response model

$$p^*(x) = 1 - \exp[-(0.00145x + 0.00013x^2)] \\ \text{(MLE risk model for mice).}$$

This model implies that, at very low doses,

$$p^*(x) = 0.00145x \\ \text{(MLE low-dose excess risk).}$$

Step 3. Uncertainty Analysis: Calculating q_1^*

The quality of the MLE is often judged by uncertainty analysis. Traditionally, uncertainty analysis is based on construction of confidence intervals or joint confidence regions for the model parameters. An approximate confidence region for q may be constructed from the mathematical statistics result that, under certain technical regularity conditions, e.g., that $L(q)$ is thrice differentiable and that q^* lies in the interior of its feasible region, the distribution of minus twice the log-likelihood ratio, namely

$$-2 \log[L(q)/L(q^*)],$$

asymptotically approaches a χ^2 distribution with k degrees of freedom (6). This asymptotic result can be used to construct approximate upper confidence limits (UCLs) for the true but unknown value of q from the MLE estimate q^* . In particular, it can be used to estimate a 95% UCL for each component of q in the multistage model, including the linear term, q_1 . This provides the theoretical basis for calculating q_1^* in programs such as GLOBAL and TOXRISK. The same technique can be used to estimate approximate confidence bands for parameters in other risk models.

The q_1^* value calculated for the low-dose slope of the multistage dose-response function for benzene is

$$q_1^* = 0.00773 \text{ expected excess tumors} \\ \text{per mg/kg/day of benzene.}$$

Equivalently, the estimated 95% upper confidence limit (UCL) on excess risk at very low doses is given by the linear function

$$p(x) = 0.00773x \\ \text{(95\% UCL on low-dose excess risk).}$$

Step 4. Extrapolating Risk from Oral Gavage Data for Mice to Inhalation Hazards for Humans

To extrapolate from the oral gavage data for mice (Table 1) to risks posed by inhalation of low levels of benzene in humans, the U.S. Environmental Protection Agency (U.S. EPA) (5) selected mg-lifetimes of benzene per unit surface area as the equivalent dose metric for extrapolating risks across routes and species. They also used ppm-hr as a dose metric for risk interpolation and extrapolation within a species. The MLE for risk in any new situation (i.e., species, route, and/or time pattern of dosing) based on a risk assessment model $f(x, q)$ and MLE parameter estimate q^* assessed in a different situation is given by the risk extrapolation formula

$$p^*(x') = f(x', q^*)$$

where x' denotes the dose corresponding to the new situation. For example, to estimate from the data in Table 1 the risk to a human male from occupational inhalation exposure to benzene, one would first calculate the "equivalent" mg/kg/day, x' , received by the worker (based on inhalation rate, body weight, and so forth) and then apply the dose-response formula derived from the mouse data

$$p^*(x') = 1 - \exp[-(0.00145x' + 0.00013x'^2)].$$

Thus, the risk created in a mouse by exposure to 1 mg/kg/day of benzene for a fraction f of its lifetime is assumed to be equal to the risk created in a human by continuous inhalation of the following concentration of benzene:

$$(20 \text{ m}^3 \text{ inhaled by humans/day}) \\ (1/70 \text{ kg for a human})(3.19 \text{ mg of} \\ \text{benzene/m}^3 \text{ of air inhaled per ppm}$$

of benzene in the inhaled air) $(1/f \text{ relative duration of lifetime human exposure})(70 \text{ kg per human/mouse weight in kg})^{1/3}$.

Using this factor for interspecies and inter-route dose extrapolation yielded a 95% UCL unit risk estimate for humans of

$$0.119 \text{ excess tumors per ppm-lifetime of exposure to benzene (extrapolated 95\% UCL).}$$

The corresponding MLE for human risk due to inhalation of benzene was 0.0161 expected excess tumors per ppm-lifetime of exposure (5).

Critique of Traditional Benzene Risk Assessment and Need for a New Approach

The preceding benzene risk assessment has three limitations.

First, the assumption of an AUC dose metric for benzene is not realistic. Many experiments have shown that different time patterns of benzene dose administration with the same AUC produce very different profiles of benzene metabolites (2) and very different hematotoxic effects (7). For example, 150 ppm of benzene inhaled for 2 hr causes more than 7 times greater production of prephenylmercapturic acid (an indicator of benzene metabolism) than 50 ppm of benzene inhaled for 6 hr (2). Similarly, exposing male CD-1 mice via inhalation to 10 ppm of benzene for 6 hr/day, 5 days/week for 10 weeks (3000 ppm-hr of exposure) has no detectable impact on marrow cellularity or on colony-forming unit, granulocyte-macrophage (CFU-GM) stem cells in bone marrow, but 100 ppm for 6 hr/day for 5 days (also 3000 ppm-hr of exposure) significantly depresses marrow colony-forming unit, spleen (CFU-S) and CFU-GM cells (7). Even more strikingly, while male NMRI mice continuously exposed to 21 ppm of benzene via inhalation for about a week show very significantly depressed marrow cellularity (cells/tibia) and CFU-GM content per tibia, mice exposed to up to 14 ppm for up to 8 weeks—a much larger AUC dose—show no significant changes in bone marrow cellularity or CFU-GM content (8). Other AUC dose violations and anomalies, such as the fact that exposure for 3 days per week may have a larger impact on erythropoiesis than exposure for 5 days per week (9), have been documented for

benzene metabolism, cytotoxicity, and genotoxicity. In light of these findings, the AUC dose metric cannot be considered biologically realistic for benzene.

The second limitation is that the q_1^* methodology may not produce predictively useful results for benzene. Reasons include the following: *a)* The regularity conditions needed to establish the asymptotic result may not hold. For example, the usual regularity conditions require that the true parameter vector q should lie in the interior of its set of possible values (so that the constraint $q \geq 0$ is not binding). Yet, the MLE for q_0 is 0. *b)* Asymptotic convergence provides no guarantees for real data set. Indeed, Monte-Carlo simulation shows that convergence is poor for some realistic-size data sets, while the nonnegativity constraints on model coefficients may lead to multi-modal distributions of q^* values and to unstable MLE estimates of risk (10). The asymptotic χ -square distribution does not always provide a useful approximation to the actual distribution of q around its MLE estimate, as evaluated by Monte-Carlo analysis. *c)* Model uncertainty is ignored. The multi-stage model does not accurately describe some real data sets (11). This possibility is ignored in analyses of uncertainty using the q_1^* methodology. Yet, such model uncertainty (e.g., about whether the chosen parametric family of risk models contains the true dose-response relation) may overwhelm uncertainty due to statistical sampling variability in determining total uncertainty about model-based risk predictions.

These limitations imply that the computed value of q_1^* may contain very little information about the true plausible upper limit on the low-dose slope. A different uncertainty analysis methodology is needed to produce more useful results. Advances in uncertainty analysis methods, including the widespread use of Monte-Carlo uncertainty analysis and other computer-intensive statistical methods (12) now offer practical approaches for improving on earlier uncertainty analyses for benzene.

The third limitation on benzene risk assessment is that the human risk predictions made by the traditional risk assessment model based on animal data do not appear to agree with epidemiological data. For example, the preceding risk model predicts that the risk to a human exposed to x ppm-years of benzene via inhalation should exceed

$$1 - \exp(-0.016x) \\ \text{(extrapolated lower bound} \\ \text{dose-response model)}.$$

(This is a lower bound because higher order terms are neglected.) Thus, a worker exposed to an average of 25 ppm-years per year for 40 years should have a lifetime tumor probability due to benzene exposure of at least

$$1 - \exp[-(0.016)(-1000 \text{ ppm-years}/70 \text{ years per lifetime})] = 0.20.$$

This prediction appears to be incompatible with observed tumor rates in highly exposed human populations. For example, Turkish shoe and handbag workers exposed to an average of over 1000 ppm-years of benzene have a lifetime excess tumor probability that appears to be less than 0.01 (13), in contrast to the lower bound prediction from the above model. The 95% UCL unit risk estimates would tend to overpredict true risks even more.

In summary, benzene risk assessment can be improved first by using a model of benzene risks that better reflects relevant benzene biology. The AUC dose metric is no longer plausible in light of current knowledge of benzene biology. The second way benzene risk assessment can be improved is by including model uncertainty in the uncertainty analysis. The possibility of model error must be addressed to accurately quantify benzene risks and uncertainties about them.

Both of these improvements are discussed in the following sections. The intended result is a set of benzene risk models that better agree with available human data and that more accurately predict human risks at low levels of benzene exposure.

Methods For Improving Benzene Risk Assessment

The following main technical ideas and methods may be used to construct new benzene risk models.

Causal Decomposition of the Cancer Risk Process. Instead of trying to quantify the relation between dose and response probability directly, it is useful to decompose the causal relation between exposure history and tumor probability into biologically meaningful causal links called "micro-relations" to quantify these links, and then to estimate the full dose-response relation by composing its constituent microrelations. As an example, the dose-response relation for benzene can be described as the composition of two microrelations, one linking administered dose to internal dose

of benzene metabolites, the other linking internal dose to tumor probability.

Model-free Curve Fitting. Instead of selecting the multistage model or another specific statistical risk model as the only model to be considered, it is potentially more realistic to express the cumulative tumor hazard as an unknown function of dose. If it is sufficiently smooth, this unknown function may be expanded as a convergent series (i.e., expressed as a weighted sum of orthonormal basis functions), and the coefficients of the first few terms may be estimated from data. This provides a data-driven, model-free approximation to the true but unknown function (14). A similar but less sophisticated approach is to expand the unknown function around zero as a MacLaurin power series. This has the advantage for risk analysis purposes that the power series model may then be interpreted as a multi-stage model with unconstrained coefficients, i.e., coefficients not constrained to be nonnegative. Leaving the coefficients unconstrained may be more realistic than constraining them to be nonnegative if the effects of exposure on cell proliferation and cytotoxicity are considered.

Maximum Entropy Bayesian Monte-Carlo Uncertainty Analysis. Instead of the q_1^* methodology's asymptotic approach to uncertainty analysis, the following approach may be taken: *a)* Let the prior probability distributions for the true but unknown response probability in each dose group be uniformly distributed between 0 and 1 (corresponding to a minimum of a priori assumptions). Use Bayes' Rule to condition this prior distribution on the observed data, i.e., on the number of animals exposed and the number responding in each dose group. The result is a posterior distribution for the true but unknown response probability in each dose group. *b)* Sample many times from these distributions (Monte-Carlo sampling) and pass a model-free dose-response curve through each set of sampled tumor probability values. In our implementation of this idea, polynomial regression was used to estimate the coefficients of the first few terms in the power series expansion of the cumulative tumor hazard function, considered as an unknown function of dose. The result is a Monte-Carlo distribution of the entire dose-response curve. If there were no uncertainty about the true response probabilities, they would uniquely determine a single dose-response curve, namely, the polynomial of lowest order that passes

through them. Because the true response probabilities are not completely determined by experimental data due to sampling variability, a probability distribution of possible dose-response curves results. This distribution directly determines the upper 95% limit on the low-dose slope of the dose-response function, as well as a wealth of other information useful for characterizing uncertainty about the dose-response function.

Biologically Based Computer Simulation Models. The causal decomposition framework suggests that the age-specific hazard rate for cancer (AML) induction by benzene depends on the administered dose history through two sets of variables: the internal dose of benzene metabolites (the output of a PBPK model) on the one hand, and the product of hematotoxic and genotoxic effects (if any) on the other. A nonlinear feedback control model is introduced to quantify the potential effects of benzene exposure on the granulocyte-macrophage (GM) lineage, specifically among proliferating cells (e.g., CFU-GM cells) that may be involved in acute myeloid leukemia (AML) induction. Although this model has not yet been validated for benzene, its predictions agree with empirical data in important respects, providing potential insights into why AUC dose metrics are inappropriate and how risk may vary with the time pattern of benzene administration. Model results suggest that no single definition or measure of dose can adequately summarize the contributions to benzene-induced hematorotoxicity and AML risk from different detailed histories of benzene administration. In general, the dose-time-response relation for benzene may be too complicated for a dose metric to exist. Instead, dynamic simulation models must be used to calculate the effects of different administered dose histories. Details of these methods are discussed next.

Maximum-entropy Bayesian Monte-Carlo Uncertainty Analysis

Suppose that prior to looking at the data, we wish to make only a "minimum-commitment" assumption about the true value of the response probability $p(x)$ associated with dose x . This may be done by assigning to each $p(x)$ a uniform probability density so that, a priori, $p(x)$ is considered equally likely to have any of its possible values (between 0 and 1) (15). Although stronger prior assumptions can be introduced by placing a priori constraints on the

maximum-entropy joint prior density of the $p(x)$ values—for example, by requiring the $p(x)$ values to increase with x —such constraints may express requirements that, against a researcher's expectations, do not hold in biological reality. In the absence of definitive biological knowledge, it seems safer to avoid introducing such a priori assumptions, instead letting the shape of the $p(x)$ curve be learned directly from the data. This may be done by conditioning the prior probability density function on the experimental data using Bayes' Rule. Suppose the experimental data consist of the set

$$D = \{x, n(x), r(x) \text{ for } M \text{ values of } x\} \\ \text{(experimental data)}$$

where M is the number of dose groups, then the revised a posteriori probability density for $p(x)$ based on the observed data point $[n(x), r(x)]$ (and assuming the maximum-entropy, i.e., uniform, prior density $U[0, 1]$ with no assumed constraints interrelating the $p(x)$ values for different values of x) is a beta distribution with parameters $n(x)r(x) + 1$ and $[1 - r(x)]n(x) + 1$. This has the mean value

$$E[p(x)|D] = E[p(x)|n(x), r(x)] \\ = [r(x)n(x) + 1] / [n(x) + 2] \\ \text{[posterior mean of } p(x)|D].$$

The beta posterior distribution for $p(x)$ thus has a mean value that is close to its observed sample value of $r(x)$ when $n(x)$ is large; moreover, the variance of the beta distribution is a decreasing function of $n(x)$ (16).

Once the posterior distribution of each $p(x)$ value has been determined, M -vectors of possible response probabilities (one for each of the M dose groups) may be randomly sampled from these distributions. Each such M -vector determines a corresponding dose-response curve, via any of several curve-fitting techniques. For illustration, we will use the relatively simplistic techniques (14) of polynomial regression. The mathematical details are as follows: The tumor probability for an animal exposed to dose x may be expressed without loss of generality as:

$$p(x) = 1 - \exp[-b(x)],$$

where $b(x)$ is the cumulative tumor hazard associated with dose x . This is a mathematical identity following from the definition of cumulative hazard. Next, the unknown

cumulative hazard function is approximated by the first few terms of a power series expansion around zero as:

$$b(x) = q_0 + q_1x + q_2x^2 + \dots + q_{M-1}x^{M-1}$$

where M is the number of dose groups. For simplicity, the following exposition assumes that $M = 4$, as in Table 1. (The formulas hold for all $M \geq 2$.) The vector of cumulative hazards, denoted by

$$b = (b_1, b_2, b_3, b_4),$$

is determined from the values of the response probabilities (sampled from their appropriate β distributions) via the inverse transformation

$$b_j = -\ln[1 - p_j],$$

where p_j denotes the sampled value of the tumor probability for dose group j . This operation may be summarized by the equation

$$b = -\ln(1 - p),$$

if it is understood that the transform is applied component-by-component to the elements of p , the vector of response probabilities, to obtain the corresponding components of b . (This is consistent with the notation in several popular mathematical computing software packages, although it does not agree with standard vector algebra notation.) The coefficient vector

$$q = (q_0, q_1, q_2, q_3)$$

can now be solved for from b and the different doses used in the experiment. The system of linear equations

$$\begin{aligned} b_1 &= q_0 + q_1x_1 + q_2x_1^2 + q_3x_1^3 \\ b_2 &= q_0 + q_1x_2 + q_2x_2^2 + q_3x_2^3 \\ b_3 &= q_0 + q_1x_3 + q_2x_3^2 + q_3x_3^3 \\ b_4 &= q_0 + q_1x_4 + q_2x_4^2 + q_3x_4^3 \end{aligned}$$

can be expressed compactly in vector-matrix notation as

$$b = qX$$

where X is the matrix with $(1, x_j, x_j^2, x_j^3)^T$ as its j th column. (The equation $b = qX$ holds also for $M > 4$ dose groups, with the quantities b , q , and X extended in the obvious way.) X is known and b can be

estimated from the observed tumor incidence rates. The corresponding value of q is determined as follows:

$$q = bX^{-1}.$$

The matrix inverse X^{-1} can be found numerically; alternatively, the special structure of X can be used to develop a simple, closed-form expression for the ij th element of its inverse.

Repeating the sampling of response probability M -vectors many times gives a Monte-Carlo sampling distribution of dose-response curves. Each sampled set of tumor probabilities, p , undergoes the following sequence of transformations:

$$p \rightarrow -\ln(1-p) = h \rightarrow hX^{-1} = q \rightarrow p(x) = 1 - \exp[-(q_0 + q_1x + q_2x^2 + q_3x^3)].$$

Thus, each p determines a corresponding dose-response curve expressed as a function of x at the right end of this sequence. Therefore, the beta distributions of the components of p imply a corresponding probability distribution for dose-response curves. This distribution of curves reflects the posterior uncertainty, after conditioning on the experimental data, about the true but unknown $p(x)$ values. Figure 1 shows the data flow diagram for the entire algorithm.

The Monte-Carlo distribution of dose-response curves allows a much richer characterization of uncertainties than does q_1^* analysis. Not only can the value of the low-dose slope be determined such that 95% of all of the sampled dose-response curves have smaller slopes—the analog to the q_1^* value in traditional analysis—but other detailed questions about the probable shape of the dose-response function can also be answered. For example, the probability distribution of the dose level at which a given level of risk (such as 1×10^{-6}) is first exceeded can be determined from the Monte-Carlo sample of dose-response curves. The conditional probability distribution of the risk at one dose level given assumptions about the risk at other dose levels can also be determined.

Methods for a Biologically Based Causal Decomposition of Dose-Response Models

The causal relation between the administered dose history and the probability that a tumor develops by any specified time is thought to be mediated by several biological phenomena, including benzene

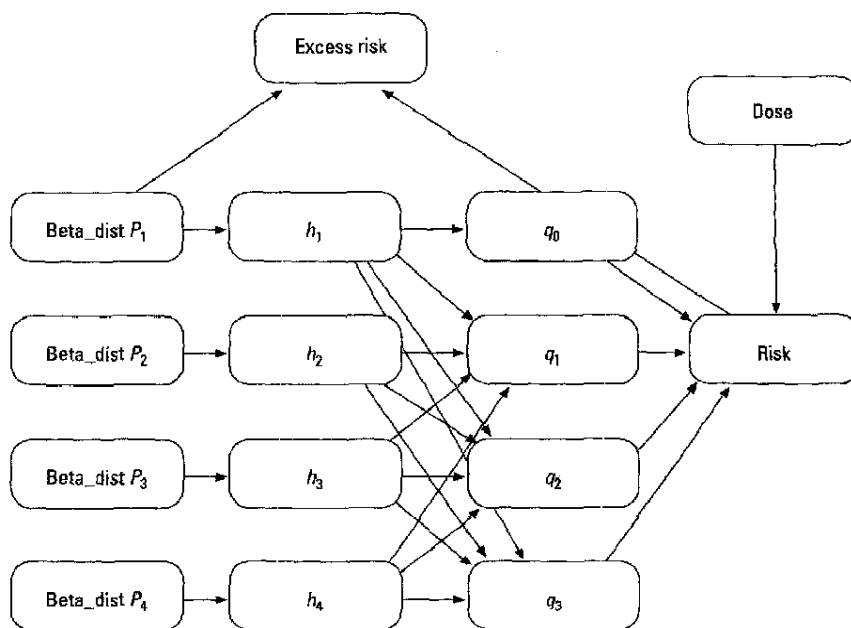


Figure 1. Data flow for the uncertainty analysis. Probability distributions for the true response probability in each dose group (denoted by P_j for group j) are derived by conditioning maximum entropy (uniform) priors on observed experimental data, leading to beta distributions. [The cumulative density functions (CDFs), representing these beta distributions are shown in Figure 2.] The CDF for each response probability is transformed to a corresponding CDF for a cumulative hazard rate via the identity $h_j = -\ln(1 - P_j)$. The vector of cumulative hazard rates is used to solve for the vector of coefficients in a polynomial regression model via the formula $q = X^{-1}h$. Finally, the coefficients in q are used to compute the risk for each of several doses, x , via the formula $p(x) = 1 - \exp[-h(x)] = 1 - \exp[-(q_0 + q_1x + q_2x^2 + \dots + q_nx^n)]$.

metabolism and pharmacokinetics, benzene-induced cytotoxicity and genotoxicity in the bone marrow, and perhaps effects on the immune system. These phenomena are typically easier to observe experimentally than tumor risk itself. To discuss how they may be used to improve benzene risk assessment, it is useful to introduce the notation

$(a \parallel b)$ = time history of quantity a
determined by the time history
of quantity b .

This notation is intended to imply that a and b are time-varying quantities, with the history of a up to and including any moment t being completely determined by the history of b up to the same moment. Other constants or parameters may have to be specified for the history $\{b(\tau), 0 \leq \tau \leq t\}$ to uniquely determine the history of $\{a(\tau), 0 \leq \tau \leq t\}$ for every value of t such that $0 \leq t \leq T$, where 0 is the time of the exposed animal's birth and T is the time of its death. If so, these additional quantities may be listed after b following the "causal conditional" sign \parallel in the above notation.

For example, the dose-time-response relation between benzene exposure and AML risk may be denoted by

$$(p \parallel x)$$

to indicate that the probability of tumor by any time t is determined by the history of benzene exposures up to time t . Another suggested notation for a closely related concept is

$$\{x\} \rightarrow \{p\},$$

read as "the history of x determines the history of p ." By contrast, $\{p \parallel x\}$ might be read as "the history of p that is determined by the history of x ." The curly brackets enclose individual time-varying quantities, or histories; thus, for example, $\{x\}$ is an abbreviation of the more explicit notation

$$\{x(t), 0 \leq t \leq T\}.$$

Now, suppose that AML risk depends on administered benzene only because of the formation of one or more benzene metabolites. Letting $\{y\}$ denote the history

(i.e., the time course) of the vector of metabolites in different physiological compartments over time, the causal situation may be diagrammed as

$$\{x\} \rightarrow \{y\} \rightarrow \{p\}.$$

In other words, $\{x\}$ determines $\{y\}$ and $\{y\}$ determines $\{p\}$. Thus, the administered dose–time–response relation ($p \parallel x$) may be decomposed as

$$(\{x\} \rightarrow \{y\}) * (\{y\} \rightarrow \{p\}),$$

where $*$ is the composition operator for composing consecutive mappings. This decomposition, based on internal dose of metabolites, may be expressed in the equivalent form

$$(p \parallel x) = (p \parallel y) * (y \parallel x).$$

The component $(y \parallel x)$ corresponds to the input–output relation of a physiologically based pharmacokinetics (PBPK) model, i.e., it maps administered dose histories $\{x\}$ into resulting time courses $\{y\}$ of benzene and its metabolites in different physiological compartments. The microrelation $(p \parallel y)$ represents an internal dose–response function.

According to recent biological theories of carcinogenesis, most cancer risks can be expressed in terms of the following microrelations:

- a PBPK model, represented by the link $\{x\} \rightarrow \{y\}$;
- a cytotoxicity and cell kinetics model, $\{y\} \rightarrow \{N\}$;
- a genotoxicity and cytogenetics effect model, $\{y\} \rightarrow \{\mu\}$;
- a model of cell initiation, $\{N, \mu\} \rightarrow \{I\}$;
- a hazard rate model, $\{I, \mu\} \rightarrow \{b\}$;
- a time-to-tumor model, $\{b\} \rightarrow \{p\}$.

In these links, the symbols are interpreted for benzene-induced AML as follows: $\{x\}$ = administered dose history; $\{y\}$ = metabolite history, giving the time courses of concentrations of benzene and benzene metabolites in different physiological compartments (e.g., blood, bone marrow, fat, muscle, richly perfused tissue, poorly perfused tissue, and various organ groups); $\{N\}$ = time course of hematopoietic cell population sizes; $\{\mu\}$ = history of cell transition rate parameters (including birth, death, differentiation, and carcinogenic transformation rates); $\{I\}$ = time course of a distinguished “initiated” (pre-malignant) cell population for AML; $\{b\}$ = cumulative

hazard function, giving the age-specific or time-specific cumulative hazard for AML induction; and $\{p\}$ = history of tumor probability as a function of time. Vector quantities are indicated in boldface. Notation of the form $\{a, b\} \rightarrow \{c\}$ means that the history $\{c\}$ is determined by the conjunction of histories $\{a\}$ and $\{b\}$. Note that, as it is defined here, the quantity $p(t)$ is greater than the probability of a detectable neoplasm being initiated by time t . To be detected, two further things must occur after a malignant cell forms: a) It must escape immune system surveillance and control; and b) it must progress to a detectable size.

A Two-stage Stochastic Model of Cancer Induction

The complete causal model for cancer induction in a two-stage stochastic model of carcinogenesis is composed from the preceding microrelations according to the following diagram:

$$\{x\} \rightarrow \{y\} \rightarrow \{\mu, N\} \rightarrow \{p_2, I\} \rightarrow \{b\} \rightarrow \{p\}.$$

Here, the initiated population, I , is interpreted as a stochastic process (a birth–death process with immigration from the “normal” stem cell compartment), and only its probability law is determined by the data $\{\mu, N\}$. The links in this causal chain may be quantified with the help of the following biomathematical model equations

$$(\{b\} \rightarrow \{p\}) \quad [1]$$

This link is quantified via the mathematical identity

$$p(t) = 1 - \exp[-b(t)].$$

Here, $p(t)$ is the probability that at least one malignant cell is initiated by time t . Interpretively, if a randomly occurring event (formation of a malignant cell) occurs at a constant average rate of α times per year, then the probability that it has not occurred after t years is $\exp(-\alpha t)$. The probability that it occurs at least once by time t is therefore $1 - \exp(-\alpha t)$. The above identity generalizes this to the case of time-varying occurrence rates, $\alpha(t)$. The cumulative hazard $b(t)$ is the integrated value of $\alpha(t)$ (its AUC) up through time t . If α is constant, then $b(t) = \alpha t$.

$$\{\mu_2, I\} \rightarrow \{b\} \quad [2]$$

This link is quantified via the formula

$$db(t)/dt = \alpha(t) = \mu_2(t)I(t),$$

where $\mu_2(t)$ is the transformation rate from initiated to malignant cells (per initiated cell per unit time) at time t . This formula has the common-sense interpretation that the expected number of malignant cells being formed at time t , namely $\alpha(t)$, is the product of the number of cells at risk of malignant transformation, $I(t)$, and the rate at which transformations occur, $\mu_2(t)$.

$$\{\mu, N\} \rightarrow \{\mu_2, I\} \quad [3]$$

This link is quantified via a stochastic differential equation with

$$dE[I(t)]/dt = (b(t) - d(t) - \mu_2(t))I(t) + \mu_1(t)N_1(t).$$

Here, $b(t)$ and $d(t)$ are birth and death rates and $\mu_1(t)$ and $\mu_2(t)$ are transformation rates (all per cell per unit time); all four are components of the parameter vector $\mu(t)$. Similarly, $N_1(t)$ refers to a component of $N(t)$ (e.g., early myeloid hematopoietic progenitor cells).

$$\{y\} \rightarrow \{\mu, N\} \quad [4]$$

This link describes the cytotoxic, cytogenetic, and genotoxic effects of benzene metabolites on different cell populations. At present, sufficient *in vivo* experimental data and biological knowledge to reliably quantify this link are not available. Instead, therefore, the link

$$\{x\} \rightarrow \{\mu, N\}$$

will be quantified using curve-fitting techniques and experimental data.

$$\{x\} \rightarrow \{y\} \quad [5]$$

This link can be quantified for benzene using any of several published PBPK models of the dynamics of benzene metabolism, distribution, and elimination (17,18).

If each link can be quantified, then the links can be composed to obtain the total dose–time–response relation

$$\begin{aligned} (\{x\} \rightarrow \{p\}) &= (\{x\} \rightarrow \{y\}) * (\{y\} \rightarrow \{\mu, N\}) \\ &\quad * (\{\mu, N\} \rightarrow \{\mu_2, I\}) \\ &\quad * (\{\mu_2, I\} \rightarrow \{b\}) * (\{b\} \rightarrow \{p\}). \end{aligned}$$

This may be written more compactly as:

$$(p||x) = (p||b) * (b||\mu_2, I) * (\mu_2, I||\mu, N) * (\mu, N||y) * (y||x).$$

However, this is only one of many possible decompositions. It is not the most useful one to use with existing benzene data since the component $(\mu, N||y)$ is highly uncertain. Two alternative decompositions that can better use available data are

$$(p||x) = (p||y) * (y||x) \\ = (p||b) * (b||\mu_2, I) * (\mu_2, I||\mu, N) * (\mu, N||x).$$

Different decompositions of the dose-time-response relation suggest different statistical strategies for quantifying it. For example,

$(p||x)$, the undecomposed relation, may be estimated by fitting curves or models to experimental data on administered doses and resulting tumors. This is the traditional approach illustrated already for the data in Table 1 using the U.S. EPA data (5) risk assessment as an example.

$(p||y) * (y||x)$ involves quantifying both the PBPK component, $(y||x)$, based on experimental data on administered doses and resulting metabolite levels and also the internal dose-time-response relation, $(p||y)$, based on experimental data on metabolite levels and tumor rates.

$(p||b) * (b||\mu_2, I) * (\mu_2, I||\mu, N) * (\mu, N||x)$ involves quantifying the relation $(\mu, N||x)$ based on experimental data on administered doses and cell level responses (cell proliferation rates, population sizes, cytotoxic impacts, and cytogenetic and genotoxic transformation rates). For the remaining components, both $(p||b)$ and $(b||\mu_2, I)$ are obtained from the previously described mathematical equations. $(\mu_2, I||\mu, N)$ involves identifying the cell population corresponding to I and the transformation corresponding to μ_2 .

Each of these strategies for decomposing and quantifying risk makes use of different data and has its own advantages. Each will be used subsequently in developing results on benzene risks.

Quantifying Microrelations by Simulating Biological Processes: PBPK and Cytotoxicity Models for Benzene

To quantify a relation such as $(y||x)$, it is usually necessary to build a dynamic

simulation model that can accept time series histories $\{x\}$ as inputs and calculate resulting output histories such as $\{y\}$. Such models typically take the form of systems of nonlinear ordinary or partial differential equations. In the simplest cases, the output history can be calculated from the input history by iterating the following discrete-time approximation for small values of the time step, s :

$$y(i+s) = y(i) + g[x(i), y(i)]s \\ \text{(Euler's method of numerical integration)}$$

where the instantaneous rate of change of y at time t is given by the formula

$$dy/dt = g[x(t), y(t)].$$

The transition rate function g , indicated in bold because it is a vector function, determines the time evolution of y from its own current value and the concurrent value of the input, x . In practice, g is obtained by biomathematical modeling of the specific biological system involved.

Two examples of links for which dynamic models have been developed for benzene are PBPK models and cytotoxicity models.

PBPK Models for Benzene

The microrelation $(y||x)$ is quantified by PBPK models of benzene metabolism and distribution. Several such models have been published and are available for use (1, 17, 18). For risk assessment purposes, we prefer the human and animal PBPK models of Travis et al. (17). These models quantify total benzene metabolites and do not track circulating metabolites. On the positive side, they have been extensively compared to and validated with human and animal experimental data. The details of the PBPK models are discussed by Travis et al. (17).

Henderson et al. (19) present the following three empirical $y(x)$ data points for mice exposed to benzene via oral gavage, as in Table 1: $y(12) = 10$; $y(40) = 25$; and $y(150) = 50$. In each of these $y(x)$ data points, the x value is the administered dose measured in mg of benzene per kg of animal body weight and the y value is the resulting internal dose, defined as the total amount of metabolites formed in mg divided by the mouse body weight in kg. The internal doses $y(x)$ corresponding to the administered dose levels of $x = 0, 25, 50$, and 100 mg/kg/day used in the NTP mouse bioassay experiment in Table 1 have not been directly measured. However, linear interpolation among these three

published data points gives interpolated values of $y^*(25) = 18.25$; $y^*(50) = 30$; and $y^*(100) = 46$. Alternatively, we can follow Bailer and Hoel (6) in assuming that the $\{x, y(x)\}$ points fall approximately on a nonlinear regression curve having the Michaelis-Menten form $y(x) = Vx/(K + x)$ where $m = (V, K)$ is a parameter vector to be estimated from data. Least-squares estimation using the preceding three data points gives $K^* = 80.75$, $V^* = 76.4$, corresponding to $y^*(25) = 18.1$, $y^*(50) = 29.2$, and $y^*(100) = 42$. These values are slightly less than the linearly interpolated values of $y^*(25) = 18.25$; $y^*(50) = 30$; and $y^*(100) = 46$. The least-squares regression procedure smoothes out sampling variability and measurement error to the extent that the underlying Michaelis-Menten model is correct. Given that the regression and interpolation values are close and that the former are more conservative, the Michaelis-Menten regression estimates could be used in the rest of the analysis.

An alternative, more biologically based modeling approach is to use a benzene PBPK model to simulate the pharmacokinetic and metabolic processes that convert administered doses to internal doses. Current PBPK models accomplish this simulation by representing the flow of benzene from compartment j to compartment i at time t (for nonmetabolizing compartments) by the equation

$$f_{ij}(t) = k_{ij}[C_j(t) - C_i(t)/P_i].$$

The rate of change in the concentration of benzene in compartment i at time t is thus

$$dC_i(t)/dt = \sum_j [f_{ij}(t) - f_{ji}(t)],$$

the flow into i from j minus the flow out of i into j . (If j is not adjacent to i , the flow between them is zero.) If benzene is metabolized in compartment i , such as liver or bone marrow, then $dC_i(t)/dt$ also includes a term $[-V_{max}C_i(t)/P_i]/[K_m + C_i(t)/P_i]$, where V_{max} and K_m are the Michaelis-Menten parameters for the aqueous component of the compartment. Given the values of the pharmacokinetic parameters k_{ij} and P_i and the metabolic parameters V_{max} and K_m , this system of differential equations can easily be integrated numerically to solve for the time series $\{C_i(t)\}$. In current PBPK models for benzene, the metabolic parameters V_{max} and k_m are highly uncertain; they have been estimated by Medinsky et al. (18) and Travis et al. (17) by seeking values that would make

model outputs match available data as well as possible.

The blood:tissue and blood:air partition coefficients P_i for benzene are available in the literature. Medinsky et al. (18) provide values of blood:tissue and blood:air partition coefficients P_i for i = liver, fat, poorly perfused tissues, richly perfused tissues, and air. Travis et al. (17) identified ranges of P_i values from the literature for i = fat, liver, muscle, organs (corresponding to richly perfused tissues) and air and then chose values within these ranges that caused the PBPK model predictions to best fit available data. The flow rate parameters k_{ij} are determined by PBPK theory as follows:

- If the flow is from j = arterial blood to i = a tissue compartment, then $k_{ij} = Q_j/V_i$, the blood flow rate (in liters per minute) through compartment i divided by the volume (in liters) of compartment i . Values of the physiological constants Q_i and V_i have been estimated for rats, mice, humans, and other species (20).
- If the flow is from j = alveolar air to i = arterial blood (or vice versa), then the value of k_{ij} is assumed to be so large that the flow $f_{ij}(t) = k_{ij}[C_j(t) - C_i(t)/N]$ can be modeled as adjusting to equilibrium instantaneously. (N denotes the blood:air partition coefficient.) Instead of having to estimate k_{ij} , therefore, the equilibrium condition $C_{alb}(t) = C_{art}(t)/N$ is assumed.
- The total flow of benzene into mixed venous blood from tissues per unit time is the sum over all tissue groups, i , of $Q_i C_i(t)/P_i$, while the total flow of blood into the venous side of the circulation per unit time is just Q , the circulation flux (l/min), which is equal to the total cardiac output rate. If it is assumed that the volume of the venous system is negligible, then the concentration of benzene in the venous blood will be the sum over i of $(Q_i/Q) C_i(t)/P_i$, i.e., it will be the average concentration of benzene in blood entering the mixed venous blood. This determines the (approximate) value of $C_{ven}(t)$, the concentration of benzene in mixed venous blood at time t .

In addition to these flow rate and equilibration assumptions, the following flow balance constraint is also assumed:

$$C_{art}(t)(Q_{alb}/N + Q) = [Q_{alb}C_{inh}(t) + \sum_i Q_i C_i(t)/P_i],$$

implying that

$$C_{art}(t) = [Q_{alb}C_{inh}(t) + \sum_i Q_i C_i(t)/P_i] / (Q_{alb}/N + Q).$$

This equates the total inflow of benzene per unit time into the alveolar-air-and-arterial-blood compartment to the total outflow of benzene per unit time from that compartment.

These assumptions and parameter estimates allow the time series of benzene concentrations in each compartment to be determined from the inhalation exposure time series $\{C_{inh}(t)\}$. If the dose administration route is oral gavage or injection (sc or ip), the PBPK model must be extended to allow for transfer of the administered bolus dose into the rest of the system. For oral gavage, the simplest expedient is to assume a first-order gastric absorption constant (transfer rate) from the GI tract to the liver. Both Medinsky et al. (18) and Travis et al. (17) make this assumption and both estimate similar rate constants in mice—0.032/min according to Travis et al. and 0.05/min according to Medinsky et al. [However, the gastrointestinal transfer rate constant estimated for rats is 0.0042 according to Medinsky et al. (18) and 0.01 according to Travis et al. (17).]

We have reimplemented the Travis et al. PBPK model for benzene (17), incorporating the preceding modeling assumptions and parameter values. This model has been extensively compared against both human and animal data on time courses of benzene concentrations in expired air, blood, and bone marrow during and following exposures to benzene via inhalation, oral gavage, and injection. It provides a useful fit to nearly all the available data sets, as discussed in detail by Travis et al. (17). The resulting simulation model for metabolism of oral gavage doses in mice predicts values for internal doses slightly greater than the ones from the Michaelis-Menten regression model. Table 2 compares the predictions from the PBPK and Michaelis-Menten internal dose models.

The Michaelis-Menten model in Table 2 was fit to the three empirical data points from which the parameter values $K^* = 80.75$ and $V^* = 76.4$ were estimated. A better fit to the PBPK model output is given by the revised Michaelis-Menten model $y^*(x) = 77.5x/(74.33 + x)$. This curve provides a compact approximation of the PBPK results. At the National Toxicology Program (NTP) dose levels of $x = 25, 50$, and

75 mg/kg, it predicts internal doses of $y^*(25) = 19.5$, $y^*(50) = 31.2$, and $y^*(100) = 45.5$, compared to the PBPK model predictions of $y^*(25) = 19.0$, $y^*(50) = 30.1$, and $y^*(100) = 44.6$. Agreement at other points is even closer. Thus, $y^*(x) = 77.5x/(74.33 + x)$ provides a useful analytic approximation to the PBPK model for predicting total metabolites formed from oral gavage doses.

Cytotoxicity and Cell Kinetics Models for Benzene

The link $(N||x)$ may be quantified by biomathematical simulation models of granulopoiesis and hematoxicity. Such models have been previously developed and tested against experimental data in mice for benzene (21) and in dogs for the immunosuppressive and myelotoxic agent cyclophosphamide (CP) (22). Although such models are relatively immature and exploratory compared to PBPK models, they do synthesize large amounts of relevant biological experimental data and may prove useful in future efforts to improve benzene dose-time-response modeling.

A computer simulation model of benzene-induced hematotoxicity is currently being developed. While it would be premature to rely on it as a basis for benzene risk assessment, it offers potential insights that may be useful enough to justify a brief outline of the model's main features. It is based primarily on the model of Steinbach et al. (22) updated with benzene cytotoxicity parameters from Scheding et al. (21). We have validated the model with published human clinical data for CP (23). The main model consists of a set of compartments representing cell populations, linked by flow equations with flow rate parameters that change over time in response to changes in the sizes of the cell populations. Several nonlinear feedback control laws operate to restore the cell population sizes to their stable equilibrium

Table 2. Mouse internal doses of benzene metabolites (mg/kg) produced from different oral gavage doses, as predicted by various models.

Dose level x	Observed values	Predicted $y(x)$ values	
		Michaelis-Menten	PBPK
1	—	0.93	1
5	—	4.5	4.9
10	—	8.4	9.2
12	10	9.9	10.7
25	—	18.1	19.0
40	25	25.3	26.2
50	—	29.2	30.1
100	—	42.3	44.6
150	50	49.7	54.5

values following any small perturbation, e.g., due to inhalation of benzene. These feedback control laws are not powerful enough to maintain or restore equilibrium cell population levels if sufficiently intense exposure continues, however. Even after cessation of dosing, the system may not return to equilibrium for many weeks.

The compartments of the cell kinetics and hematotoxicity model are defined and abbreviated as follows:

- **S** = early myeloid stem cells and hematopoietic progenitor cells (HPCs). These cells are further subdivided into two subpopulations:
 - cycling_HPCs = actively cycling HPCs (i.e., those participating in the mitotic cycle);
 - resting_HPCs = resting or dormant HPCs (not actively cycling).
- **CFU** = granulopoietic committed stem cells. Biologically, these correspond roughly to CFU-GM cells, i.e., to early myeloid cells that can form GM colonies. To model accurately both the mean and the variance of transit times through this compartment, it is necessary to account for the age structure of the CFU-GM population. The partial differential equation (PDE) describing the time evolution of the age-structured population is approximated by dividing the compartment into 10 fictitious subcompartments representing cells of different ages. The age/maturity structure of this population also matters in determining the dynamic response of the hematopoietic system to CP-induced stresses. Subdividing it into 10 subcompartments allows shifts in its age composition to be approximated.
- **P** = proliferative pool, consisting roughly of myeloblasts, promyelocytes, and myelocytes. This aggregate compartment is also subdivided into 10 subcompartments to simulate the changing age structure of the proliferative subpopulation over time.
- **M** = maturation pool, in which cells finish maturing and lose their remaining proliferation and differentiation capacities, thus becoming no longer susceptible to carcinogenic transformations.
- **R** = bone marrow reserve of mature granulocytes. Granulocytes are released from this compartment to peripheral blood as needed to replace losses due to normal cell senescence and death or to emergencies such as bleeding.
- **B** = mature granulocytes in circulating blood. In comparing model predictions

to experimental data, this compartment represents an approximation to white blood cells (WBCs) (although lymphoid cells are not modeled).

These cell populations are the components of the vector N in the microrelation ($N||x$). Other aspects of granulopoiesis, such as the shift of some hematopoiesis to the spleen under stress, deterioration of the stromal microenvironment of the bone marrow over time, or migration of stem cells from the marrow into the peripheral blood and back are ignored in this model since their significance for the dosing scenarios considered is expected to be small.

The flow equations linking the compartments of the model (22) express three key ideas: *a*) the flux of living cells out of each compartment enters its downstream neighbor; *b*) the flux of dead cells from each proliferative compartment is proportional to the number of cells currently in the compartment and to the concentration of administered benzene weighted by a compartment-specific cytotoxicity parameter; and *c*) the magnitude of the flux out of a compartment is directly proportional to the compartment's current size (i.e., to the number of cells in it) and inversely proportional to the transit time of cells through the compartment.

The number of cells in any of the compartments at any time is found from the flow equations by incrementing the immediately previous number to reflect inflows from the upstream compartment and amplifications due to cell divisions (for the early, proliferative compartments); and by decrementing the result to reflect outflows to downstream compartments and losses due to cell death. Cell death rates due to benzene metabolites (per susceptible cell per unit time) are assumed to depend on administered concentration of benzene, as discussed by Scheduling et al. (21).

The values of key flow rate parameters are determined via nonlinear feedback control equations from the sizes of various cell populations. The hematopoietic system exerts several forms of feedback control to regulate its own behavior and compensate for cytotoxic and other stresses (21). Based on several decades of experimental data in animal models, the following five specific quantities are controlled by feedback control loops in the model.

HPC_recruitment_rate = fractional recruitment rate of resting HPCs, defined as the fraction of resting HPCs recruited into active cycling per unit time. This rate is controlled not only by the concurrent

numbers of resting and cycling stem cells, but also by the sizes of the downstream cell populations in the CFU-GM and subsequent compartments up through and including the bone marrow reserve. If the more mature cell populations depart from their steady-state levels in response to hematopoietic stresses due to dosing, the recruitment and production of stem cells counter-adjusts to help move the system back toward its equilibrium levels of these populations.

HPC_differentiation_fraction = The fraction of HPCs that differentiate (rather than self-renewing). A simplified model of the cell cycle is used for stem cells, indicating the fraction (A/S) that are actively dividing at each moment. As each cell completes the mitotic cycle, it either differentiates, making a transition into the first CFU-GM proliferative compartment, or it passes immediately into G1, starting the next round of cell division, or it lapses into G0, the resting state. The fraction that differentiates is determined by the current numbers of resting and cycling stem cells.

CFU_birth_rate = fractional birth rate of CFU-GM cells (average birth rate per CFU-GM cell per unit time). If the more mature GM lineage downstream from these GM-committed stem cells becomes depleted, the division (i.e., birth) rate of the CFU-GM cells increases to compensate. Conversely, if the downstream cell populations are well stocked, then the birth rate among CFU-GM cells slows. The CFU-GM birth rate is also affected by feed-forward from the earlier stem cell compartment.

P_birth_rate = fractional birth rate of proliferative cells. These cells are assumed to respond only to signals from CFU-GM and later cells, up through and including the bone marrow reserve, but not to signals from the earlier (pre-CFU-GM) stem cells.

Release_rate_to_blood = fractional release rate of mature granulocytes from bone marrow reserve to peripheral blood (per reserve cell per unit time). This rate is controlled by a direct feedback loop from the peripheral blood compartment; there is no longer range control from upstream cell populations.

The nonlinear feedback control equations determining these quantities are summarized in the appendix. The specific functional forms used are empirical approximations obtained by fitting smooth curves to experimental data and to extreme or limiting conditions (e.g., maximal birth rates obtained under experimental conditions of maximum stimulation). Sensitivity analysis reveals that the exact forms of the

equations are less important than their overall shapes. Steinbach et al. (22) discuss in greater detail the experimental basis and the assumptions used to derive the feedback control equations.

Using the Decomposition Framework for Risk Extrapolation

A principle advantage of the causal decomposition approach to dose-response modeling is that it provides an alternative to dose metrics as a basis for risk modeling. Thus, even though a predictively useful dose metric may not exist for benzene, risks can still be quantified and extrapolated. The principle of extrapolation is illustrated by the following formula:

$$(p \parallel x)_{\text{human, inhalation}} = (p \parallel y)_{\text{mouse, gavage}} * (y \parallel x)_{\text{human, inhalation}}$$

This formula shows that the relation $(p \parallel y)$ estimated from experimental data involving mice exposed via oral gavage can be composed with the relation $(y \parallel x)$ estimated from data (or from a PBPK model) for humans exposed via inhalation. The result is an extrapolated (across species and route) estimate of the dose-time-response relation $(p \parallel x)$. Each term in the decomposition is subscripted to indicate the type of data from which it has been estimated. Ideally, all terms would be estimated from the same type of data (e.g., sex, age, species, strain, route of administration, etc.), namely, the data type to which the risk model as a whole is to be applied. In practice, however, if data for humans exposed via inhalation (for example) are not available to quantify some of the components, then animal data may be substituted. This provides a method of incorporating human data where it is available and using only animal data where necessary. Models quantified in this way can subsequently be refined when more relevant data become available.

Results

Results of Monte-Carlo Uncertainty Analysis: The Slope of the Dose-Response Curve Is Unlikely to Be Positive at the Origin

Figures 2 to 6 show the results of applying the maximum-entropy Bayesian Monte-Carlo uncertainty analysis technique to the data in Table 1. The posterior cumulative distribution functions (CDFs) for the tumor response probabilities after conditioning on

the sample sizes and observed tumor incidence rates are shown in Figure 2. These provide the input to the rest of the computations summarized in Figure 1. Figure 3 shows the CDF for q_1 . Notice that there is a high probability (about 90%) that q_1 is negative, suggesting that there is no excess risk at low doses. Figure 4 shows this explicitly: as administered dose increases from 1 mg/kg/day to 5 mg/kg/day, the PDF for risk shifts left. However, at 25 mg/kg/day, the lowest dose group for which data are available in Table 1, the risk is significantly increased compared to the zero-dose risk. The interpretation of negative risk values at lower doses is discussed below. Figure 5 shows the expected value of the dose-response function (the solid curve in the middle) with 75 and 95% probability bands around it. These bands are to be interpreted in terms of the Monte-Carlo-derived posterior PDF or CDF for the response probability at each dose level rather than in terms of classical hypothesis testing. Thus, they replace the 75 and 95% upper and lower confidence limits that would be derived from the asymptotic χ^2 distribution using a traditional q_1^* approach. Figure 6 shows analogous results for excess risk, defined as total risk minus background (zero-dose) risk. The 95% upper probability band for excess risk has a positive slope of 0.004 and is linear at low doses. This is about half the value of q_1^* derived using the traditional (GLOBAL '82) approach.

All these results are based on simple polynomial regression as a curve-fitting technique, as described in the methods section. More sophisticated nonparametric smoothing (e.g., splines) could be used instead.

Results of Traditional Multistage Modeling Using PBPK Internal Doses: The Dose-Response Curve Has Zero Slope at the Origin

The strategy of decomposing the benzene dose-response relation as

$$(p \parallel x) = (p \parallel y) * (y \parallel x)$$

has been implemented for the data in Table 1 (25). The internal dose microrrelation $(y \parallel x)$ for mice was quantified in three steps: *a*) The PBPK model of Travis et al. (17) was used to calculate the steady-state amounts of total benzene metabolites produced per day from different amounts of benzene administered per day. *b*) A nonlinear (Michaelis-Menten) regression model was used to quantify the relation between x , the administered daily dose of

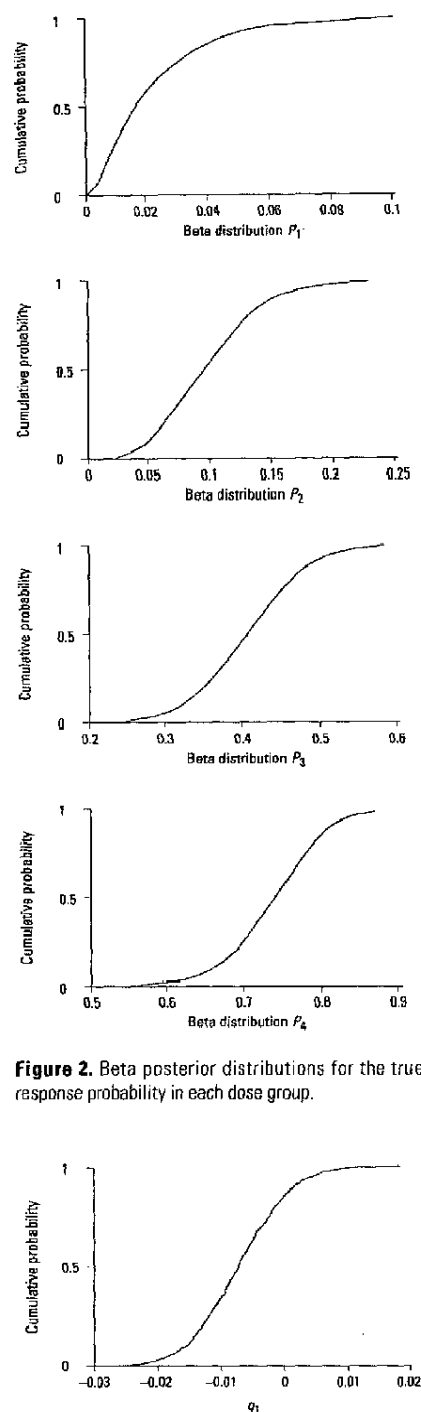


Figure 2. Beta posterior distributions for the true response probability in each dose group.

Figure 3. Cumulative probability density function (CDF) for q_1 . The CDF is obtained via Monte-Carlo simulation, by propagating numbers sampled from the beta distributions on the left of Figure 1 through the rest of the diagram. The CDF for each derived quantity in Figure 1 is obtained via an algebraic combination of the (sampled) values of the quantities pointing into it.

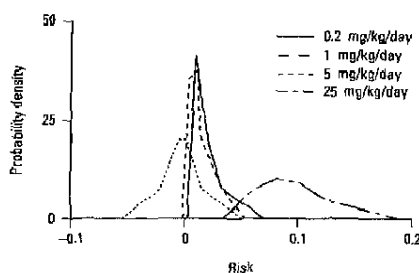


Figure 4. Sensitivity analysis of the probability density function of risk to different administered concentrations. At doses $x=0.2$ and $x=1$ mg/kg/day, the probability mass is concentrated near zero (and is less than 0.1). For $x=5$ mg/kg/day, the probability mass is shifted left: the probability of a positive tumor response is less than at lower doses. At $x=25$ mg/kg/day, however, it is very likely that the risk is positive.

benzene, and y , the total amount of benzene metabolites formed per day. c) The fractions of total benzene metabolites formed along different specific metabolic paths (e.g., leading to hydroquinone conjugates, muconic acid, etc.) were quantified based on empirical data.

The mouse PBPK model relation between administered dose and internal dose (defined as total benzene metabolites formed per day mg/kg/day in mice exposed via oral gavage to x mg/kg/day of benzene) was well approximated by the following nonlinear (Michaelis-Menten) regression model:

$$(y|x)_{\text{mouse, gavage}} = 76.4x / (80.75 + x) \\ (x \text{ in mg/kg/day via gavage})$$

The corresponding relation for humans exposed to benzene inhalation derived from a benzene PBPK model for humans (17) was

$$(y|x)_{\text{human, inhalation}} = 9.525x / (132.2 + x) \\ (x \text{ in ppm via inhalation}).$$

Other choices of route and dose regimen lead to similar formulas. Each such formula gives a useful numerical approximation of the results of the full dynamic PBPK model for a specific choice of the exposure summary variable, x , the metabolism variable, y , and for a range of administered dose scenarios.

Next, the microrelation $(p||y)$ was quantified using a traditional multistage dose-response model, taking both total benzene metabolites and the amounts of

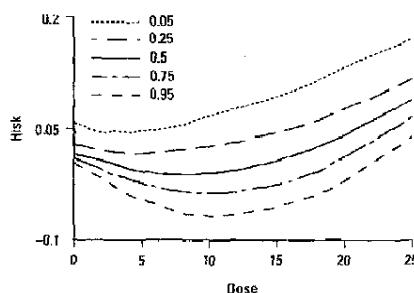


Figure 5. Probability bands for the dose-response relation. For doses below about 10 mg/kg/day, it is unlikely that there is a positive excess tumor risk due to benzene exposure.

metabolites formed along specific metabolic pathways (corresponding to specific components of y) as the dose variable, and using the tumor data in Table 1 for response data. The internal dose versus observed response data points for total benzene metabolites are shown in Table 3.

The projected internal dose levels (the y^* values) in Table 3 are estimated from the corresponding administered dose levels ($x = 0, 25, 50$, and 75 mg/kg/day) using the Michaelis-Menten statistical interpolation model $y^*(x) = 76.4x / (80.75 + x)$. The traditional multistage methodology (implemented in GLOBAL '82) applied to the internal dose estimates in Table 3 produced the following results:

The MLE cancer probability from y mg/kg/day of benzene metabolites (for male B6C3F1 mice) is

$$(p||y) = 1 - \exp(-0.000019y^3) \\ (\text{MLE internal dose-response} \\ \text{model for mice}).$$

The estimated 95% UCL for the low-dose slope, calculated via the q_1^* methodology, is:

$$q_1^* = 0.00554. \\ (\text{estimated 95\% UCL}).$$

Table 3. Estimated internal dose-response functions for mice.

Administered dose x : mg/kg/day	0	25	50	75
Estimated dose, y^* (mg/kg/day)	0	18	29	42
% of animals with squamous cell carcinomas (50 animals/dose group)	0	8	40	74

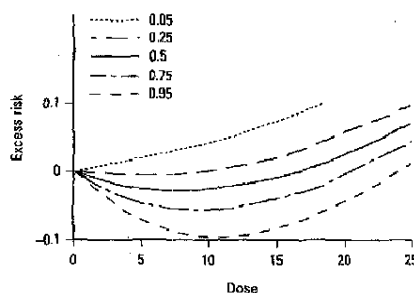


Figure 6. Probability bands for excess risk. The expected excess risk (solid line) is negative at doses below 15 mg/kg/day. However, there is a 5% (but not a 25%) chance that it is positive even at very low doses, as predicted by the q_1^* methodology.

For humans, the dose-response function is estimated as follows:

$$(p||x)_{\text{human, inhalation}} = (p||y)_{\text{human, inhalation}} \\ * (y||x)_{\text{human, inhalation}} \\ = (p||y)_{\text{human, inhalation}} \\ * 9.525x / (132.2 + x).$$

In the absence of specific data from which to estimate $(p||y)_{\text{human, inhalation}}$, the internal dose-response function estimated for mice (the most sensitive known species, sex, and strain) was used instead:

$$\text{estimated } (p||x)_{\text{human, inhalation}} \\ = (p||y)_{\text{mouse, gavage}} * 9.525x / (132.2 + x) \\ = 1 - \exp[-0.000019(9.525x / (132.2 + x))^3] \\ (\text{estimated human dose-response function}).$$

Note that the MLE model for $(p||y)_{\text{mouse, gavage}}$ has a slope of zero as y approaches zero (and hence as x approaches zero). The estimated total dose-response relation $(p||x)_{\text{human, inhalation}}$ is nonlinear with a slope of zero at $x = 0$. A maximum entropy Bayesian Monte-Carlo uncertainty analysis of the mouse internal dose-response function $(p||y)_{\text{mouse, gavage}}$ confirmed that there is no increase in expected risk at doses below about 14 mg/kg/day.

Applications of the Human Risk Model to Different Dosing Scenarios

Tables 4 and 5 show the results of applying the foregoing risk model to two human exposure scenarios of practical interest: a daily dosing scenario (Table 4) and a continuous exposure scenario (Table 5). In Table 5, the risk to humans from routine exposure is shown as a function of an unknown constant, k . This is defined as the ratio of the steady-state, biologically effective dose

Table 4. Estimated human dose-response model for 8 hr/day lifetime inhalation exposures to benzene.

	Concentration, ppm						
	1	5	10	25	50	100	200
Average daily internal dose (mg/kg/day) from PBPK model	0.067	0.32	0.63	1.5	2.6	4.2	6.0
Lifetime excess cancer risk, based on mouse dose-response function	6×10^{-9}	6×10^{-7}	5×10^{-6}	6×10^{-5}	3×10^{-4}	0.0014	0.004

Table 5. Individual excess cancer risks from lifetime inhalation exposure to 1 ppm of benzene, for different values of k (= mouse:human internal dose ratio).

Dose measure	Risk model	Value of k		
		0.01	0.1	1
Administered dose	$1.45 \times 10^{-3} x + 1.31 \times 10^{-4} x^2$	0.0016	0.0016	0.0016
Total metabolites	$1.86 \times 10^{-5} y^3$	$1.1 \text{ E-}13$	$1.1 \text{ E-}10$	$1.1 \text{ E-}7$
Total HQC dose	$1.51 \times 10^{-3} y^3$	$0.8 \text{ E-}13$	$0.8 \text{ E-}10$	$0.8 \text{ E-}7$
Total MA dose	$1.51 \times 10^{-2} y^3$	$8.1 \text{ E-}13$	$8.1 \text{ E-}10$	$8.1 \text{ E-}7$
Total metabolites, two-stage LMS	$6.21 \times 10^{-4} y^2$	$2.0 \text{ E-}9$	$2.0 \text{ E-}7$	$2.0 \text{ E-}5$

Abbreviations: HQC, hydroquinone conjugates; MA, muconic acid; LMS, linearized multistage risk model.

Table 6. Maximum likelihood linearized multistage risk models fit to male mouse squamous cell cancer incidence data for different surrogate internal dose measures and tumor end points.

y = Dose surrogate	Maximum likelihood LMS risk model	Description
End point—all squamous cell carcinomas		
Administered dose	$1.45 \times 10^{-3} x + 1.31 \times 10^{-4} x^2$	Linear-quadratic
Total metabolites	$1.86 \times 10^{-5} y^3$	Cubic
Total HQC dose	$1.51 \times 10^{-3} y^3$	Cubic
Total MA dose	$1.51 \times 10^{-2} y^3$	Cubic
AUC HQC dose	$4.14 y^2 + 2.44 y^3$	Quadratic-cubic
AUC MA dose	$0.448 y^2 + 0.004 y^3$	Quadratic-cubic
End point—squamous cell carcinomas of Zymbal and preputial glands ^a		
Zymbal gland SCC	$6.13 \times 10^{-6} y^3$	Cubic
Preputial gland SCC, total metabolites	$1.51 \times 10^{-4} y^2 + 8.89 \times 10^{-6} y^3$	Quadratic-cubic

Abbreviations: LMS, linear multistage risk models; SCC, squamous cell carcinoma. ^aSpecific tumor type end points should only be analyzed by time-to-tumor models that adjust for competing risks due to other tumors. These quantal models are for purposes of illustration only. From Cox and Ricci (20).

of the carcinogenic metabolite in humans to its value in mice, assuming low, constant administered concentrations of benzene. Table 6 shows that the main finding of a zero slope for the internal dose-response function at the origin (i.e., low-dose non-linearity) holds for many choices of internal doses and end point. Details of the computations summarized in Table 6, are presented by Cox and Ricci (20).

Results of Computer Simulation Modeling: the Hematotoxicity Dose-Time-Response Relation Has a Negative Slope at Low Concentrations

Ideally, the results of risk calculations based on the decomposition

$$(p||x) = (p||y) * (y||x)$$

would be cross-checked and improved using the results of alternative decompositions such as

$$(p||x) = (p||N, \mu) * (N, \mu||x).$$

Although some data are available to quantify the cytotoxic effects on N and the cytotoxic and genotoxic effects on μ of inhalation exposure to benzene (7), a detailed calculation of benzene dose-response relations by this decomposition analogous to the one based on PBPK modeling of internal doses, has not been done. Nonetheless, it is now possible to provide a key component of such a model by using a simulation model of benzene-induced hematotoxicity. Figures 7 to 9 show results from the biologically based computer simulation model

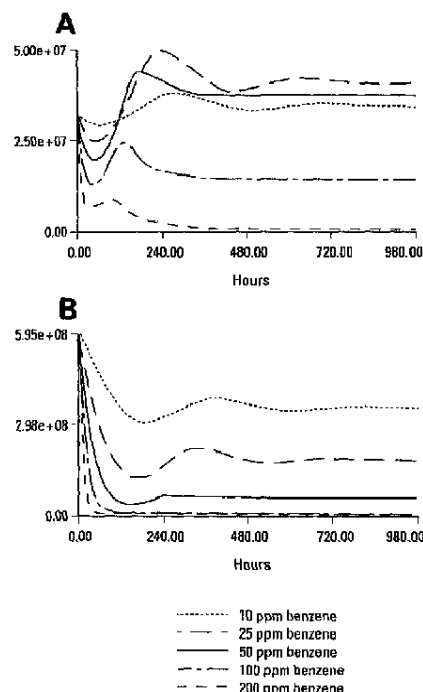


Figure 7. Computer simulations of early and late CFU-GM responses to different benzene concentrations. Curves show the simulated time courses of CFU-GM cells—(A) earliest CFU-GM 1; (B) most mature CFU-GM 10—in response to constant exposures to (approximately) 10, 25, 50, 100, and 200 ppm of benzene, respectively.

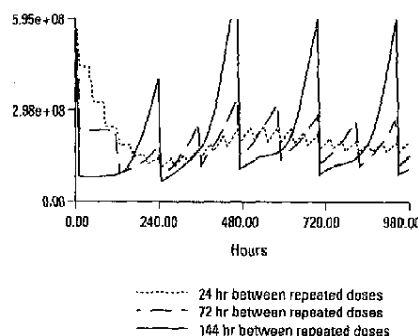


Figure 8. Different amounts of spacing between repeated 8-hr inhalation doses of benzene cause very different hematopoietic responses for the same administered dose (AUC) per unit time. Curves correspond to 24, 72, and 144 hr between repeated doses.

used to quantify the relation $(N||x)$. There are two major conclusions from this model. a) The function $N(x) = (N||x)$, expressing the steady-state size of early myeloid proliferative cell population (e.g., HPCs or CFU-GM) as a function of the administered benzene concentration

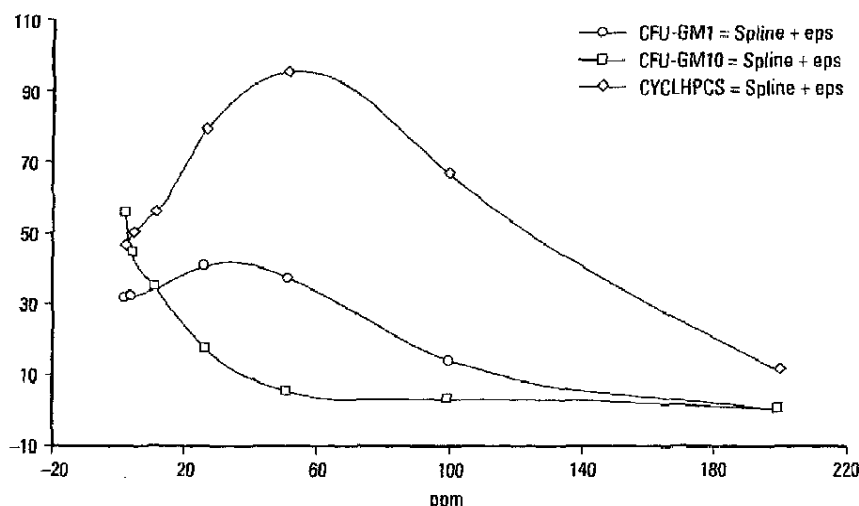


Figure 9. Steady-state relations between benzene exposure and hematopoietic cell population sizes predicted by a dynamic simulation model. Benzene dose axis may need to be rescaled for humans.

assuming a constant exposure scenario, is a nonmonotonic function of benzene concentration (Figure 9). At sufficiently low concentrations, benzene exposure is predicted to increase cell population sizes (for early HPCs and immature CFU-GM cells, but not for very mature CFU-GM cells). In Figures 7 to 9, CFU-GM1 is the earliest of the 10 age-structured subcompartments of the CFU-GM population, while CFU-GM10 is the last of these subcompartments. Figure 9 plots the steady-state levels in Figure 7 against administered concentration. It shows that the relation between ppm of benzene and predicted steady-state values of N (where N is one of the early CFU-GM or HPC populations) is positive at sufficiently low concentrations but negative at higher concentrations. If it is conjectured that AML risk is increased only by exposure scenarios that significantly depress CFU-GM and/or HPC populations (e.g., because recruitment into active cycling or clonal expansion of pre-leukemic stem cells becomes more likely), then the predicted shape of the $(N||x)$ relation would be consistent with the hypothesis that benzene at sufficiently low concentrations does not increase (and might even decrease) AML risk. *b*) The same AUC of administered dose may have profoundly different effects on hematotoxicity depending on how it is administered over time. For the same AUC per unit time of administered benzene, higher concentrations may lead to disproportionately large hematotoxic effects (Figure 8). The

nominal administered concentration in Figure 8 is 100 ppm. Different concentrations with the same time pattern of administration produce similar results for a broad range of administered concentrations. Since the benzene hematotoxic potency factors in the model were originally developed for mice rather than for humans, however, human data are needed to calibrate the model more exactly for benzene.

Although the simulation model produces a wealth of other results, it should be developed and validated further before being used for risk assessment. Nonetheless, the model-based conclusions appear to be consistent with the results arrived at by other methods.

Discussion

Bayesian Monte-Carlo Uncertainty Analysis Results

The 95% upper probability bands on low-dose slopes for excess risk derived by Bayesian Monte-Carlo uncertainty analysis (Figure 6) are of the same order of magnitude as the q_1^* values derived by traditional methods. However, the Monte-Carlo uncertainty analysis in Figures 3 to 6 also reveals a great deal of additional information that many decision makers—including all who follow the expected utility paradigm of rational decision making (23)—would consider relevant. Perhaps most striking is the fact that the expected value of the excess risk is negative for administered doses smaller than about

15 mg/kg/day (Figure 6). Taken at face value, this finding suggests that the health protection benefits from reducing doses below this level are very uncertain, and may not be significantly different from zero. Notice that while the total risk of tumor defined as a probability or hazard rate, must be nonnegative, the risk due to a specific cause such as benzene can perfectly well be negative. For example, risk might be reduced if benzene exposures in the 1 to 15 mg/g/day range were to suppress the expression of carcinogenic damage (e.g., by killing or preventing the division of affected cells) or were to reduce the sizes of the populations at risk of carcinogenic damage. Since benzene was once used as a human chemotherapeutic agent, such effects on cell kinetics would not be completely unexpected. Without speculating about such mechanisms, the interpretation of the negative risk values in Figures 4 through 6 is that the expected lifetime tumor probability at low doses is less than the expected lifetime tumor probability at zero dose.

Traditional Risk Model Based on Internal Doses

The benzene risk model based on PBPK modeling and use of internal doses, namely,

$$(p||x)_{\text{human, inhalation}} = 1 - \exp\{-0.000019[9.525x/(132.2 + x)]^3\},$$

differs significantly at low dose from the traditional (administered dose) benzene risk model discussed in the introduction. For example, at $x = 12$ mg/kg/day, mice form approximately $y = 76.4 * 12 / (80.75 + 12) = 10$ mg/kg/day of total benzene metabolites, giving a projected lifetime cancer risk of $P^*(10) = 0.019$ based on the internal dose model. By contrast, the administered dose risk model $p^*(x) = 1 - \exp[-(0.00145x + 0.00013x^2)]$ gives a predicted lifetime cancer risk level of $p^*(12) = 0.035$ based on the administered dose model. Thus, fitting the dose-response function to internal doses instead of administered doses changes the maximum likelihood linearized multistage model from linear-quadratic to cubic, reducing estimated risk by about a factor of 2 at an administered dose level of $x = 12$ mg/kg/day. [This oral gavage dose level corresponds roughly to a 6-hr inhalation dose of about 10 ppm based on total metabolites formed (19).] Similar calculations for $x = 1$ mg/kg/day give $p^*(0.93) = 0.000015$ for the Michaelis-Menten model [or $p^*(1) = 0.000019$ for

the PBPK model] compared to $p^*(1) = 0.0016$ for the administered dose model, i.e., at an administered dose level of 1 mg/kg, risk estimated from the internal dose model is about 1% of the risk estimated from the administered dose model. At lower doses (corresponding to inhalation doses below 1 ppm) the cubic internal dose model gives lifetime cancer risk predictions less than 1% as large as those from the linear-quadratic administered dose model.

There are several reasons to believe that the cubic internal dose-response risk model may be more realistic than the administered dose-response model. If a Weibull model of the form $p(y) = 1 - \exp[-(a + by^c)]$ is fit to the four internal dose response points in Table 4, the result is the MLE Weibull risk model

$$(p||y) = 1 - \exp(-0.000015y^{3.065}).$$

Thus, the conclusion that the relation between internal dose and probability of response is approximately cubic (rather than linear or quadratic) does not appear to be only an artifact of using the multi-stage family of risk models. Also, the internal dose risk model provides a slightly better fit than the administered dose risk model to the available animal dose-response data. The sample values of $(p||x)$ at $x = 25, 50$, and 100 mg/kg/day are $(p||x) = 0.08, 0.40$, and 0.74 . The corresponding values predicted by the administered dose linear-quadratic model are $0.11, 0.33$, and 0.77 , while the values predicted by the cubic model with the Michaelis-Menten internal dose model are $0.105, 0.37$, and 0.755 . In each case, the best-fitting (cubic) internal dose-response model predicts values of $(p||x)$ slightly closer to the observed ones than the corresponding predictions from the best fitting (linear-quadratic) administered dose-response model. Quantitatively, the internal dose model provides a better fit to the data, with the maximized value of the log-likelihood function being -76.53 for the internal dose model compared to -77.15 for the administered dose model. Finally, the internal dose-response model appears to be much more consistent with epidemiological data on tumors in heavily benzene-exposed workers (Turkish shoe workers and Chinese workers) than the administered dose-response model.

For an administered dose of $x = 1$ mg/kg/day, the internal dose model predicts an MLE risk of only 0.0000155 , which is about 1% as large as the MLE risk predicted by the administered dose model. On the other hand, the respective upper UCLs for risk at $x = 1$ mg/kg/day are 0.005 for the internal dose model compared to 0.008 for the administered dose model. Thus, even though the MLE estimates differ by a factor of 100 , the 95% UCLs based on traditional q_1^* uncertainty analysis are very similar. This insensitivity of 95% UCLs to significant variations in risk models is typical; it has been both criticized as a defect and praised as a virtue of the q_1^* approach.

Biologically Based Risk Assessment Modeling

Biologically based computer simulation models are available for quantifying some of the key microrelations thought to be involved in benzene leukemia causation, including $(y||x)$ and $(N||x)$. Where such models exist, they can be used to quantify the output histories, like $\{y\}$ and $\{N\}$, caused by any benzene administered dose history, $\{x\}$. Instead of using dose metrics for risk extrapolation, therefore, it is possible to construct explicit computer simulation models of the pharmacokinetic and pharmacodynamic processes (ideally including hematotoxicity and genotoxicity) mediating between benzene exposure and AML induction.

Such biologically based risk assessment (BBRA) models have the advantage of making explicit, testable predictions of observable quantities (such as quantities of benzene metabolites formed, cell population sizes, and so forth) as well as tumor risk, which is much harder to observe directly. They also bypass the need to rely upon dose metrics, which may not exist in any useful form for benzene. On the other hand, BBRA models are often too complex and detailed and require too many uncertain inputs to be used for regulatory risk assessment and public health risk management.

A principle expounded in this article to avoid complexity is that if administered dose histories are restricted to simple forms, then the full dynamic simulation approach becomes unnecessary. Specifically, under a wide range of conditions, a constant administered dose eventually produces a constant steady-state output for

each component of $\{y\}$ and $\{N\}$. In this case, letting x denote the constant applied concentration (in ppm, for inhalation exposures) and letting $y(x)$ and $N(x)$ denote the resulting steady-state levels of a specified benzene metabolite and cell population, respectively, it is possible to quantify the relations $(y||x)$ and $(N||x)$ using nonlinear regression models such as

$$y(x) = Vx / (K + x) \\ \text{(Michaelis-Menten nonlinear} \\ \text{regression model)}$$

where V and K are model parameters to be estimated from observed data (such as $\{x, y\}$ pairs, in this example). Alternatively, if a PBPK model is available, as is the case for benzene, it can be used to quantify many (x, y) pairs, and then a relatively simple nonlinear regression model can be used to describe the results, bypassing the need to use the PBPK model itself when only simple, constant dose scenarios are being considered.

Conclusion

The conclusions of greatest potential practical interest from the preceding analyses are as follows. *a*) There is no evidence of a positive relation between benzene exposure and tumor probability (or causal antecedents such as depressed myeloid stem cell and HPC populations) at benzene concentrations below 1 ppm. *b*) Different analytic approaches—including Bayesian Monte-Carlo uncertainty analysis and PBPK-based internal dose modeling—suggest that the $(p||x)$ curve relating benzene concentration to AML risk at sufficiently low, constant concentrations of benzene approaches a zero or negative slope as concentration falls below about 10 ppm. *c*) Empirical data and biologically based risk models agree that, for the same total administered dose (AUC), higher concentrations of benzene may cause disproportionately large hematotoxic responses.

Quantitatively, none of the statistical or biological evidence investigated suggests that reducing benzene concentrations below 1 ppm would have any detectable health benefits. The hypothesis of a zero or nonpositive slope for excess risks due to inhalation of low concentrations of benzene may have sufficiently interesting public health policy implications to merit further evaluation.

Appendix: Benzene Hematotoxicity Model Equations and Parameter Values

HPC Cells

$$d(\text{cycling_HPCs})/dt = \text{recruited_from_resting} + (\text{cell_divisions} - \text{HPC_deaths}) - \text{return_to_resting} - \text{differentiated}$$

$$d(\text{resting_HPCs})/dt = \text{return_to_resting} - \text{recruited_from_resting}$$

$$\text{recruited_from_resting} = \text{HPC_recruitment_rate} * \text{resting_HPCs}$$

$$\text{cell_divisions} = \text{recruited_from_resting} \text{ \{since each recruited cell divides at mitosis\}}$$

$$\text{HPC_deaths} = \text{cycling_HPCs} * \text{cytotoxic_potency_for_HPCs} * \text{metabolite_in_cells}$$

$$\text{metabolite_in_cells} = \text{output from PK model}$$

$$\text{return_to_resting} = (1 - \text{HPC_differentiation_fraction}) * \text{cycling_HPCs} / \text{HPC_cycle_time}$$

$$\text{differentiated} = \text{HPC_differentiation_fraction} * \text{cycling_HPCs} / \text{HPC_cycle_time}$$

CFU-GM Cells

$$d\text{CFU}(1, t)/dt = \text{differentiated} - [\text{CFU}(1, t)/\text{CFU_transit_time}] - (\text{CFU_death_rate} - \text{CFU_birth_rate}) * \text{CFU}(1, t)$$

$$d\text{CFU}(i, t)/dt = [\text{CFU}(i-1, t)/\text{CFU_transit_time}] - [\text{CFU}(i, t)/\text{CFU_transit_time}] - (\text{CFU_death_rate} - \text{CFU_birth_rate}) * \text{CFU}(i, t), \text{ for } i = 2, 3, \dots, 10$$

GM Proliferation

$$dP(1, t) = \text{inflow_from_CFU_GM}(t) - [P(1, t)/P_transit_time] - (P_death_rate - P_birth_rate) * P(1, t)$$

$$dP(i, t) = [P(i-1, t)/P_transit_time] - [P(i, t)/P_transit_time] - (P_death_rate - P_birth_rate) * P(i, t), \text{ for } i = 2, 3, \dots, 10$$

$$\text{inflow_from_CFU_GM}(t) = G(10, t)/\text{CFU_transit_time}$$

$$P_death_rate = \text{cytotoxicity_potency_for_proliferating_cells} * \text{metabolite_in_cells}$$

Nonproliferative Populations

$$dM/dt = \text{precursors_in} - \text{matured_cells_out}$$

$$\text{precursors_in} = \text{matured_cells_out}(t) = M(t)/\text{maturation_transit_time}$$

$$dR/dt = \text{matured_cells_out} - \text{release_rate_to_blood} * R$$

$$dB/dt = \text{release_rate_to_blood} * R - \text{decay_rate} * B$$

$$\text{HPC_recruitment_rate} =$$

$$k1 * \text{EXP}(-v1 * \text{resting_HPCs} - v1 * \text{cycling_HPCs}) + k2 * \text{EXP}(-v2 * Z) + k3$$

$$k1 = 0.36 \text{ \{per hr\}}$$

$$k3 = 0.008 \text{ \{per hr\}}$$

$$v1 = 1.6955\text{E-}6$$

$$k2 = 0.04 \text{ \{per hr\}}$$

$$v2 = 3.1918\text{E-}11$$

$$Z = (\text{signal_Z} + \text{CFU_GM}) \text{ \{= } 3.208\text{E}9 + 3.469\text{E}10 + 12.8\text{E}10 \text{ in steady state\}}$$

$$\text{signal_Z} = \text{proliferating_pool} + \text{maturation_pool} + \text{marrow_reserve} = P + M + R$$

$$\text{HPC_differentiation_fraction} =$$

$$1 - \text{MAX}(0, 0.5 + 1.76\text{E-}9 * (\text{steady_state_HPCs} - \text{resting_HPCs} - \text{cycling_HPCs}))$$

$$\text{steady_state_HPCs} = 3.125\text{E}8 \text{ \{HPC cells in steady state\}}$$

$$\text{CFU_birth_rate} = 0.1/(3 + \min(0.2, k4 * \text{resting_HPCs} + k4 * \text{cycling_HPCs}) + \min(0.2, k5 * \text{CFU_GM}) + \min(0.2, k6 * \text{signal_Z}))$$

$$k4 = 3.6418\text{E-}9$$

$$k6 = 6.9953\text{E-}12$$

$$(22, \text{calculated from steady-state conditions})$$

$$k5 = 3.5476\text{E-}10$$

$$\text{CFU_GM} = \text{CFU_GM}_1 + \text{CFU_GM}_2 + \text{CFU_GM}_3 + \text{CFU_GM}_4 + \text{CFU_GM}_5 + \text{CFU_GM}_6 + \text{CFU_GM}_7 + \text{CFU_GM}_8 + \text{CFU_GM}_9 + \text{CFU_GM}_{10}$$

$$\text{CFU_transit_time} = 10 \text{ \{hr for each of 10 compartments = 100 hr\}}$$

$$\text{CFU_GM_cytotoxicity_death_rate} = \text{metabolite_in_cells} * \text{cytotoxic_potency_for_CFU_GM}$$

$$\text{cytotoxic_potency_for_CFU_GM} = 0.13 \text{ \{per proliferating cell/hr\}}$$

$$P_birth_rate = 0.07/(1 + \min(0.6, k7 * Z)) \text{ (19, eq. [16])}$$

$$k7 = 3.7289666\text{E-}12; Z = (\text{signal_Z} + \text{CFU_GM})$$

$$P_transit_time = 5.6 \text{ \{hr for each of 10 compartments = 56 hr\}}$$

$$\text{cytotoxic_potency_for_proliferating_cells} = 0.13 \text{ \{per proliferating cell/hr\}}$$

$$P = \text{proliferating_pool} = P1 + P2 + P3 + P4 + P5 + P6 + P7 + P8 + P9 + P10$$

$$P_death_rate = \text{cytotoxicity_death_rate_for_proliferating_cells} = \text{metabolite_in_cells} * \text{cytotoxic_potency_for_proliferating_cells}$$

$$\text{cytotoxic_potency_for_proliferating_cells} = 0.13 \text{ \{per proliferating cell/hr\}}$$

$$\text{release_rate_to_blood} = k8 - k9 * \text{EXP}(-v3 * \text{GRF}) = 2 - 1.9875 * \text{EXP}(-6.309\text{E-}4 * \text{GRF})$$

$$\text{blood_decay_rate} = 0.1 \text{ (22)}$$

$$\text{GRF_decay_rate} = 0.1 \text{ \{per hr (22)\}}$$

REFERENCES

1. Spear RC, Bois FY, Woodruff T, Asslander D, Parker J, Selvin S. Modeling benzene pharmacokinetics across three sets of animal data: parametric sensitivity and risk implications. *Risk Anal* 11:641-654 (1991).
2. Crump KS, Allen BC. Quantitative Estimates of the Risk of Leukemia from Occupational Exposures to Benzene. Final Report to the Occupational Safety and Health Administration. Ruston, LA:Science Research Systems, 1984.
3. Occupational Safety and Health Administration (OSHA). 29 CFR Part 1910. Occupational Exposure to Benzene: Final Rule. Fed Reg September 11, 1987, 34,460-34,508.
4. U.S. EPA. Interim Quantitative Cancer Unit Risk Estimates Due to Inhalation of Benzene. EPA-600/X-85-022. Washington:U.S. Environmental Protection Agency, 1985.
5. Lawless JF. Statistical Models and Methods for Lifetime Data. New York:John Wiley & Sons, 1981.
6. Bailer AJ, Hoel DG. Metabolite-based internal doses used in a risk assessment of benzene. *Environ Health Perspect* 82:177-184, 1989.
7. Green JD, Snyder CA, LoBue J, Goldstein BD, Albert RE. Acute and chronic dose/response effect of benzene inhalation on the peripheral blood, bone marrow, and spleen cells of CD-1 male mice. *Toxicol Appl Pharmacol* 59:204-214 (1981).
8. Toft K, Olofsson T, Tunek A, Berlin M. Toxic effects of mouse bone marrow caused by inhalation of benzene. *Arch Toxicol* 51:295-302 (1982).
9. Luke CA, Tice RR, Drew RT. The effects of exposure regimen and duration on benzene-induced bone marrow damage in mice. I: Sex comparison on DBA/2 mice. *Mutat Res* 203:251-272 (1988).
10. Portier C, Hoel D. Low-dose extrapolation using the multi-stage model. *Biometrics* 39:341-352 (1983).
11. Cox LA Jr. An exact analysis of the multistage model explaining dose-response concavity. *Risk Anal* 15(3):359-368 (1995).
12. Hjorth JS. Urban. Computer Intensive Statistical Methods: Validation, Model Selection, and Bootstrap. New York:Chapman and Hall, 1994.
13. Aksoy M. Hematotoxicity and carcinogenicity of benzene. *Environ Health Perspect* 82:193-197 (1989).
14. Tarter ME, Lock MD. Model-Free Curve Estimation. New York:Chapman and Hall, 1993.
15. Jessop A. Informed Assessments: An Introduction to Information, Entropy, and Statistics. New York:Ellis Horwood, 1995.
16. DeGroot MH. Probability and Statistics. New York:Addison Wesley 1975;266-268.
17. Travis C, Quillen JL, Arms AD. Pharmacokinetics of benzene. *Toxicol Appl Pharmacol* 102:400-420 (1990).
18. Medinsky M, Sabourin PJ, Lucier G, Birnbaum LS, Henderson RF. A physiological model for simulation of benzene metabolism by rats and mice. *Toxicol Appl Pharmacol* 99:193-206 (1989).
19. Henderson RF, Sabourin PJ, Bechtold WE, Griffith WC, Medinsky MA, Birnbaum LS, Lucier GW. The effect of dose, dose rate, route of administration, and species on tissue and blood levels of benzene metabolites. *Environ Health Perspect* 82:9-17 (1989).
20. Cox LA Jr, Ricci PF. Reassessing benzene cancer risks using internal doses. *Risk Anal* 12(3):401-410 (1992).
21. Scheding S, Loeffler M, Schmitz S, Seidel HJ, Wichmann HE. Hematotoxic effects of benzene analyzed by mathematical modeling. *Toxicology* 72:265-279 (1992).
22. Steinbach KH, Raffler H, Pabst G, Fliedner TM. A mathematical model of canine granulocytopoiesis. *J Math Biol* 10:1-12 (1981).
23. Quiggin J. Generalized Expected Utility Theory: The Rank-Dependent Model. Boston:Kluwer Academic Publishers, 1993.